Published in final edited form as:

Am J Epidemiol. 2014 August 1; 180(3): 288–296. doi:10.1093/aje/kwu111.

Body Mass Index Categories in Observational Studies of Weight and Risk of Death

Katherine M. Flegal[*], **Brian K. Kit**, and **Barry I. Graubard**

National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, Maryland (Katherine M. Flegal, Brian K. Kit); and Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland (Barry I. Graubard)

Abstract

The World Health Organization (Geneva, Switzerland) and the National Heart, Lung, and Blood Institute (Bethesda, Maryland) have developed standard categories of body mass index (BMI) (calculated as weight (kg)/height (m)^2) of less than 18.5 (underweight), 18.5–24.9 (normal weight), 25.0–29.9 (overweight), and 30.0 or more (obesity). Nevertheless, studies of BMI and the risk of death sometimes use nonstandard BMI categories that vary across studies. In a meta-analysis of 8 large studies that used nonstandard BMI categories and were published between 1999 and 2014 and included 5.8 million participants, hazard ratios tended to be small throughout the range of overweight and normal weight. Risks were similar between subjects of high-normal weight (BMI of approximately 23.0–24.9) and those of low overweight (BMI of approximately 25.0–27.4). In an example using national survey data, minor variations in the reference category affected hazard ratios. For example, choosing high-normal weight (BMI of 23.0–24.9) instead of standard normal weight (BMI of 18.5–24.9) as the reference category produced higher nonsignificant hazard ratios (1.05 vs. 0.97 for men and 1.06 vs. 1.02 for women) for the standard overweight category (BMI of 25.0–29.9). Use of the standard BMI groupings avoids problems of ad hoc and post hoc category selection and facilitates between-study comparisons. The ways in which BMI data are categorized and reported may shape inferences about the degree of risk for various BMI categories.

Keywords

body mass index; body weight; epidemiologic methods; mortality; obesity; overweight

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Initiative used an iterative process of consultation and revision to develop recommendations on what should be included in an accurate and complete report of an observational study, taking into account empirical evidence and methodological considerations (1). One of the recommendations of the resultant STROBE Statement is that investigators "explain how

[*]Correspondence to Dr. Katherine M. Flegal, National Center for Health Statistics, Centers for Disease Control and Prevention, 3311 Toledo Road, Room 4336, Hyattsville, MD 20782 (kmf2@cdc.gov).

The findings and conclusions in this report are those of the authors and not necessarily the official views of the Centers for Disease Control and Prevention or the National Cancer Institute.

Conflict of interest: none declared.

quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why." (2, p.1637) In a survey of the epidemiologic literature, however, Turner et al. (3) found that, in many cases, no explanation of the choice of exposure categories was provided. They also note that, ". . .deliberate or subconscious data dredging could lead to a choice of grouping that accentuates an association thus increasing the risk of a false positive finding, and/or an exaggerated estimate of the exposure/outcome relationship." (3, p.7)

Studies of weight and risk of death commonly assess weight by using body mass index (BMI) (calculated as weight (kg)/height (m)^2), a form of weight adjusted for height. BMI is often categorized for purposes of analysis and presentation. However, there have been few discussions of which BMI categories to use and why. Standard BMI categories were developed by the World Health Organization (WHO) (4) and the National Heart, Lung, and Blood Institute (NHLBI) (5) in the 1990s, with BMI groupings of less than 18.5, 18.5–24.9, 25.0–29.9, and 30.0 or above. The NHLBI designated these as underweight, normal weight, overweight, and obesity. Nevertheless, a number of studies of BMI and risk of death since then have used a variety of nonstandard BMI categories. Studies sometimes incorporate the same cut points as the WHO/NHLBI categories but use finer divisions. For example, Adams et al. (6) used the cut points of 18.5, 25, and 30 but divided the normal weight category into 3 groups (BMI of 18.5–20.9, 21.0–23.4, and 23.5–24.9) and the overweight category into 3 groups (BMI of 25.0–26.4, 26.5–27.9, and 28.0–29.9). In our literature searches for published data on prospective studies of BMI and risk of death in adults (7), we found that, of the studies published since 2000 that used BMI categories, roughly half used the standard WHO/NHLBI categories for at least part of their analyses; the remainder used a wide variety of nonstandard BMI categories.

It can be difficult to interpret, evaluate, or summarize results when nonstandard categories are used. The objective of this paper is to discuss some aspects of using nonstandard categories of BMI, particularly within the normal weight and overweight ranges, using as examples a meta-analysis of recent large studies and an example from US national survey data.

METHODS

In the course of a previous literature search (7), we identified 13 large studies (each with more than 100,000 participants) of BMI and the risk of death that had used cut points identical to or within 0.1 of the standard cut points of 18.5, 25.0, and 30.0 but had subdivided the range of overweight and normal weight into finer BMI groupings. Of these, we selected the 7 studies (8–14) that provided hazard ratios and standard errors for no more than 4 subdivisions of normal weight and 3 of overweight. We also included a recently published study (15) that met the same criteria. The selected studies include pooled studies of US cohorts (8), European cohorts (13), East Asian cohorts (14), and Indian/Bangladeshi cohorts (14), as well as individual cohort studies from Korea (11), Austria (12), Australia (15), and the United States (9, 10). Weight and height data were self-reported in 4 studies (8–10, 15) and measured in the other studies (11–14). The full samples included 5.8 million participants and more than 582,000 deaths.

Of the 13 studies we originally identified, 2 studies (16, 17) were excluded from the summary because they used large numbers of subdivisions; 1 study (18) was excluded because it did not provide hazard ratios; and 3 studies (6, 19, 20) were excluded because the same data sets had already been included in the pooled study of US cohorts (8). The BMI groupings used in the selected studies are shown in Table 1. We added nomenclature based on the standard NHLBI categories and described groupings as low or high overweight and as low-, mid-, or high-normal weight, although the exact BMI values in those groups were not identical across studies, and the studies did not use these terms. The selected studies all used the high-normal weight category as the reference. We extracted the adjusted hazard ratios for each grouping from the published articles. We used a random-effects model (21) to summarize the results, and we based statistically significant heterogeneity (i.e., interstudy variance) on a 2-sided P value of less than 0.05.

The selected studies used different covariates in the final models and a variety of approaches, including various combinations of deletion of early deaths and deletions or adjustments for preexisting disease. We used the final analyses presented for the entire sample when available, and also the final analyses presented for never smokers, when available. Two studies (9, 15) presented results for never smokers but not for the full sample; 1 study (10) presented results for the full sample but not separately for never smokers.

To show the potential effects of different reference categories on hazard ratios for overweight and obesity, we also used as an example the National Health and Nutrition Examination Survey (NHANES) mortality data through 2006 for NHANES I, NHANES II, and NHANES III for those under 70 years of age at examination and limited to no more than 25 years of follow-up. This was simply chosen as an example to illustrate the effects of varying the reference category when there is a modest curvilinear relationship. For this analysis, we used Cox proportional hazards models with age as the time-line and adjusted for smoking status, race/ethnic group, and alcohol consumption, as previously described (22). The analytical data set included 32,294 participants with 9,380 deaths. We examined the effects of the following 5 different BMI reference categories: less than 25.0, 18.5–24.9, 20.0–24.9, 20.0–22.9, and 23.0–24.9. We estimated hazard ratios for overweight and obesity relative to each reference category in turn.

RESULTS

Full samples

The findings in the full samples are displayed in Table 2 for men and Table 3 for women. All studies had selected high-normal weight at the reference category, in most cases with no explanation. In all studies, underweight was associated with significantly higher risk of death relative to high-normal weight. With only a few exceptions, both low-normal weight and obesity were also associated with significantly higher risk of death relative to high-normal weight. However, both mid-normal weight and low overweight were generally not significantly different from high-normal weight, with hazard ratios varying slightly above and below 1. With 2 exceptions, the hazard ratios for low overweight were lower than the hazard ratios for mid-normal weight. High overweight was inconsistently associated with slightly higher risk of death relative to high-normal weight.

Never smokers

Of the 8 studies, 7 presented results separately for never smokers, with results as shown in Table 2 (for men) and Table 3 (for women). These results are based on considerably smaller samples including roughly 25% of the numbers of deaths in the full samples and, thus, they have reduced power to detect significant effects. As for the full samples, underweight, low-normal weight, high overweight, and obesity all tended to be associated with higher risk of death relative to high-normal weight. For both mid-normal weight and low overweight, hazard ratios relative to high-normal weight tended to vary slightly above and below 1, with point estimates for low overweight most often lower than estimates for mid-normal weight.

Summarized results

The summarized results are shown in Tables 2 and 3 separately by sex and separately for the full samples and for never smokers only. These results should be considered only approximations because the exact BMI values encompassed by these categories varied across studies. Nevertheless, some patterns are fairly consistent. Underweight was significantly associated with higher risk of death relative to high-normal weight. With 1 exception, the low-normal weight category was also significantly associated with higher risk of death relative to high-normal weight even in never smokers and even after extensive exclusions related to preexisting illness in several studies (8, 9, 14, 15). Hazard ratios were also elevated for the high overweight category, although not to the same degree as for low-normal weight. The low-normal weight category had higher hazard ratios than the high overweight category. The mid-normal weight category and the low overweight category tended to be similar to the high-normal weight reference category. In studies that used measured weight and height data, the low overweight category did not differ significantly from the high-normal weight category either for the full samples or for the never-smoking samples. Despite the variation among studies in BMI categories, populations studied, selection factors, adjustment factors, geographical location, and other factors, there was no statistically significant heterogeneity overall for the mid-normal weight category or the low overweight category.

Effects of varying reference categories in a data example from the NHANES

The effects on hazard ratios for overweight of varying the reference categories, using a data example from the NHANES, are shown in Table 4. Hazard ratios for overweight, grade 1 obesity, and grades 2–3 obesity are displayed in Table 4 relative to the following BMI categories: less than 25.0 (underweight and normal weight), 18.5–24.9 (normal weight), 20.0–24.9 (combined mid- and high-normal weight), 20.0–22.9 (mid-normal weight), and 23.0–24.9 (high-normal weight) by sex, for the full sample. The use of a narrower and higher reference category progressively increased the point estimates. When BMI less than 25 was used as the reference category, the hazard ratios for overweight were 0.95 for men and 0.98 for women. The narrower reference category of BMI of 18.5–24.9 produced slightly higher hazard ratios of 0.97 for men and 1.02 for women. Using the high-normal category as the reference always produced the highest hazard ratios for overweight (i.e., 1.05 for men and 1.06 for women).

DISCUSSION

Many studies of BMI and risk of death use nonstandard BMI categories that differ widely from study to study. Here, we present the results from 8 large studies with a total sample size of 5.8 million that subdivided the standard NHLBI normal weight (BMI 18.5–24.9) and overweight (BMI 25.0–29.9) ranges into subgroupings that we have termed underweight, low-normal weight, mid-normal weight, high-normal weight, low overweight, and high overweight. All of these studies chose the high-normal weight category as the reference category with little or no explanation. Within the range of normal weight in these studies, the low-normal category was associated with the highest risk of death, and the high-normal category was associated with the lowest risk of death. In most cases, there were no statistically significant differences between low overweight and high-normal weight. The low overweight category had risks similar to the mid-normal weight category and lower risks than the low-normal weight category. The high overweight category had risks that were similar to or lower than the low-normal weight category. Underweight and obesity were associated with higher risk relative to the high-normal weight category.

These results are consistent with those of other large studies that have found lower risk of death in the low overweight category than in the mid- or low-normal weight range. In the Prospective Studies Collaboration (18), all-cause mortality risk for both men and women was higher for those with BMI values of 20–22.5 than for those with BMI values of 25–27.5, and even higher for those with BMI values of 17.5–<20. A similar observation in a study of 2 million Norwegians with measured height and weight led Engelund (16) in 2003 to suggest that the “normal range” of BMI should be shifted upward because mortality rates were higher for those in the mid- and low-normal weight range than for those in the low overweight range. A large study in China that used measured height and weight (17) found a hazard ratio of 1.00 for those with BMI of 25.0–26.9 relative to a reference category of BMI of 24.0–24.9. The hazard ratios were 1.09 for those with BMI of 23.0–23.9 and 1.11 for those with BMI of 22.0–22.9, relative to those with BMI of 24.0–24.9.

None of the large studies tabulated here described a clear rationale for the choice of BMI categories. Several studies (9, 13, 14) note that combinations of these categories would correspond to the cutoff points proposed by the World Health Organization but do not give a reason why this is advantageous and do not make use of such combinations. Berrington de Gonzalez et al. (8, p. 2213) gave a rationale for their reference category on the basis of preliminary data analysis, stating that, “We defined a BMI of 22.5 to 24.9 as the referent category on the basis of a preliminary analysis indicating that this was usually the range of BMI associated with the lowest mortality.”

Other studies have also reported category choices that were based on preliminary data analyses. In 2 examples (23, 24), null estimates for overweight were not published because of preliminary results that showed no higher risk of death in the overweight category. Livingston and Ko (24, p. 18) combined the normal weight and overweight categories, stating that, "Initial observation of the data revealed that minimal mortality occurred in the BMI = 24.9–29.9 category compared with the normal range of BMI = 18.5–24.9. Thus, these two categories were combined . . ." He et al. (23, p. 1126) dropped the overweight and obesity categories from their analyses because, "As compared with normal weight (a body-mass index of 18.5 to 24.9), overweight or obesity was not associated with increased mortality." In 3 examples (8, 25, 26), a reference category was chosen on the basis of preliminary analyses showing that it would increase the hazard ratios in higher BMI categories. The examples above illustrate a type of publication bias (27), whereby the form in which results are published is affected by preliminary analysis.

The hazard ratios for comparisons of categories within the normal and overweight ranges are often extremely small, many in the range of 0.95–1.05, which Siontis and Ioannidis (28) have described as "tiny" hazard ratios. Siontis and Ioannidis point out that when effects are this small, "Cautious interpretation is warranted, since most of these effects could be eliminated with even minimal biases and their importance is uncertain" (28, p. 1292). As discussed by Ioannidis (29), the combination of flexible analyses and selective reporting can lead to wide variations in hazard ratios even within a single data set.

All of the selected studies used the high-normal weight category as the reference. The use of high-normal weight as the reference, rather than the mid-normal weight category, tends to produce a higher hazard ratio for the standard overweight category. We used a data example from the NHANES to illustrate the possible effects on the hazard ratio for overweight and obesity of the reference category, comparing the effects of BMI reference categories of less than 25.0, 18.5–24.9, 20.0–24.9, 20.0–22.9, and 23.0–24.9. For both men and women, the hazard ratios increased as the lower bound of the reference category increased. Among the categories studied, the hazard ratios were highest when the high-normal weight category (BMI of 23.0–24.9) was used as the reference category. Froslie et al. (30, p. 3) argue that the choice of reference BMI category can ". . . give different impressions to the reader" and obscure the interpretation, providing an example in which the hazard ratio in the highest BMI category more than doubled when a different reference category was used. Baik et al. (31) used a reference category of BMI of 23–24.9 for full analyses but a different reference category of BMI of less than 23 for age-specific analyses. Their abstract reported a relative risk of 1.19 for BMI values of 25–26.9 using the new reference category of BMI less than 23, but the estimate would have been 0.98 if they had used their original reference category.

Beyond the issue of which category to use as the reference, effect estimates and statistical power may also vary with the cut points chosen to delineate the categories. As pointed out by Schulgen et al. (32, p. 173), "One way of selecting a cut-point is to use the one at which the most impressive effect of the exposure variable on the outcome is observed. This approach might be called "outcome-oriented." Careful interpretation and adjustment are required to qualify the final result obtained using this strategy." Altman et al. (33) have critiqued the statistical properties of choosing a cut point to maximize the statistical significance of an association.

If the objective is to describe the shape or find the nadir of the BMI–mortality risk relationship, categories may not be the best approach. The use of categories constrains the identification of the low point. For example, if the lowest risk is in those with BMI values of 24.0–25.9, as was found by Lin et al. (34), a categorization using groups of 23.0–24.9 and 25.0–26.9 could not identify this category. The approach of creating categories and then selecting the category with the lowest hazard ratio by inspection, without statistical testing, is inadequate to deal with the statistical issues that arise (35, 36). To examine the shape of the curve without imposing categories, other approaches, such as linear splines, can be used (37, 38). Wong et al. (39) used fractional polynomials and found the nadir of the BMI–mortality risk curve in the overweight range for the average US man and in the normal weight range for the average US woman, results that were slightly different from their findings when standard BMI categories were used. Gilboa et al. (40), studying a different outcome, found that standard BMI categories were useful but that additional modeling with splines provided more insight regarding dose-response relationships within categories.

The effects on interpretation of using self-reported rather than measured weight and height data should also be considered. In the studies considered here, in comparisons of low overweight relative to high-normal weight among never smokers, studies with self-reported weight and height data showed small but significantly higher hazard ratios in contrast to studies using measured data, which showed smaller and nonsignificant results. The same phenomenon of higher hazard ratios when self-reported weight and height are used than when measured weight is used has been observed in other studies (7, 41) and is consistent with the effects predicted from the characteristic errors of self-reported weight and height (42, 43). Misclassification into the wrong BMI categories when self-reported weight and height data are used is often quite high. Spencer et al. (44) found that approximately 15% of those classified as overweight by self-report were actually obese; this will tend to increase the apparent risk in the overweight category. In addition, more than 25% of those classified as overweight by measured data were classified as normal weight by self-reported data. These high levels of misclassification suggest that self-reported data are unlikely to give accurate estimates of the risks associated with a specific BMI category. Attempts to correct self-reported weight and height data by the use of linear regression models do not eliminate systematic reporting errors (45).

Several aspects of the use of nonstandard BMI categories can lead to difficulties in interpretation. Throughout the range of overweight and normal weight, hazard ratios are small and can be affected by minor variations in the choice of categories. When there is a curvilinear relation of BMI to risk of death, the use of high-normal weight rather than normal weight as the reference category produces a higher relative risk for overweight. Use of the high-normal weight reference category with the standard overweight category obscures the similarities of low overweight and high-normal weight. The use of many different sets of BMI categories makes it difficult to summarize results across studies. Choices based on preliminary inspections of the data may introduce a form of publication bias. The use of nonstandard, ad hoc categories that differ among studies increases the apparent variability in the results. As noted elsewhere (7), the use of the predefined standard BMI groupings of underweight (BMI of <18.5), normal weight (BMI of 18.5–24.9), overweight (BMI of 25.0–29.9), and obesity (BMI of ≥30.0) as defined by the WHO and the NHLBI avoids issues of ad hoc and post hoc selection of categories and can facilitate between-study comparisons. Even in studies that also present their results using finer categories, the standard BMI groupings can be used as part of the analysis. These are not mutually exclusive procedures. The way in which BMI data are categorized and reported shapes inferences about the degree of risk associated with various BMI categories.

Acknowledgments

There was no external funding for this study.

We thank Dr. Eliseo Guallar (Johns Hopkins University) and Dr. Jochen Klenk (University of Ulm) for providing further information about their studies.

Abbreviations

BMI	body mass index
NHANES	National Health and Nutrition Examination Survey
NHLBI	National Heart, Lung and Blood Institute
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
WHO	World Health Organization

References

1. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007; 4(10):e296. [PubMed: 17941714]
2. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Med. 2007; 4(10):e297. [PubMed: 17941715]
3. Turner EL, Dobson JE, Pocock SJ. Categorisation of continuous risk factors in epidemiological publications: a survey of current practice. Epidemiol Perspect Innov. 2010; 7:9. [PubMed: 20950423]
4. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. World Health Organ Tech Rep Ser. 2000; 894:i–xii. 1–253. [PubMed: 11234459]
5. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: executive summary. Expert Panel on the Identification, Evaluation, and Treatment of Overweight in Adults. Am J Clin Nutr. 1998; 68(4):899–917. [PubMed: 9771869]
6. Adams KF, Schatzkin A, Harris TB, et al. Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. N Engl J Med. 2006; 355(8):763–778. [PubMed: 16926275]
7. Flegal KM, Kit BK, Orpana H, et al. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. JAMA. 2013; 309(1):71–82. [PubMed: 23280227]
8. Berrington de Gonzalez A, Hartge P, Cerhan JR, et al. Body-mass index and mortality among 1.46 million white adults. N Engl J Med. 2010; 363(23):2211–2219. [PubMed: 21121834]
9. Calle EE, Thun MJ, Petrelli JM, et al. Body-mass index and mortality in a prospective cohort of U.S. adults. N Engl J Med. 1999; 341(15):1097–1105. [PubMed: 10511607]
10. Jacobs EJ, Newton CC, Wang Y, et al. Waist circumference and all-cause mortality in a large US cohort. Arch Intern Med. 2010; 170(15):1293–1301. [PubMed: 20696950]
11. Jee SH, Sull JW, Park J, et al. Body-mass index and mortality in Korean men and women. N Engl J Med. 2006; 355(8):779–787. [PubMed: 16926276]
12. Klenk J, Nagel G, Ulmer H, et al. Body mass index and mortality: results of a cohort of 184,697 adults in Austria. Eur J Epidemiol. 2009; 24(2):83–91. [PubMed: 19184464]

13. Pischon T, Boeing H, Hoffmann K, et al. General and abdominal adiposity and risk of death in Europe. N Engl J Med. 2008; 359(20):2105–2120. [PubMed: 19005195]
14. Zheng W, McLerran DF, Rolland B, et al. Association between body-mass index and risk of death in more than 1 million Asians. N Engl J Med. 2011; 364(8):719–729. [PubMed: 21345101]
15. Joshy G, Korda RJ, Bauman A, et al. Investigation of methodological factors potentially underlying the apparently paradoxical findings on body mass index and all-cause mortality. PLoS One. 2014; 9(2):e88641. [PubMed: 24533128]
16. Engeland A, Bjørge T, Selmer RM, et al. Height and body mass index in relation to total mortality. Epidemiology. 2003; 14(3):293–299. [PubMed: 12859029]
17. Gu D, He J, Duan X, et al. Body weight and mortality among men and women in China. JAMA. 2006; 295(7):776–783. [PubMed: 16478900]
18. Whitlock G, Lewington S, et al. Prospective Studies Collaboration. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. Lancet. 2009; 373(9669):1083–1096. [PubMed: 19299006]
19. Hu FB, Willett WC, Li T, et al. Adiposity as compared with physical activity in predicting mortality among women. N Engl J Med. 2004; 351(26):2694–2703. [PubMed: 15616204]
20. Koster A, Leitzmann MF, Schatzkin A, et al. The combined relations of adiposity and smoking on mortality. Am J Clin Nutr. 2008; 88(5):1206–1212. [PubMed: 18996854]
21. DerSimonian R, Kacker R. Random-effects model for meta-analysis of clinical trials: an update. Contemp Clin Trials. 2007; 28(2):105–114. [PubMed: 16807131]
22. Flegal KM, Graubard BI, Williamson DF, et al. Excess deaths associated with underweight, overweight, and obesity. JAMA. 2005; 293(15):1861–1867. [PubMed: 15840860]
23. He J, Gu D, Wu X, et al. Major causes of death among men and women in China. N Engl J Med. 2005; 353(11):1124–1134. [PubMed: 16162883]
24. Livingston EH, Ko CY. Effect of diabetes and hypertension on obesity-related mortality. Surgery. 2005; 137(1):16–25. [PubMed: 15614276]
25. Berraho M, Nejjari C, Raherison C, et al. Body mass index, disability, and 13-year mortality in older French adults. J Aging Health. 2010; 22(1):68–83. [PubMed: 19920206]
26. Kulminski AM, Arbeev KG, Kulminskaya IV, et al. Body mass index and nine-year mortality in disabled and nondisabled older U.S. individuals. J Am Geriatr Soc. 2008; 56(1):105–110. [PubMed: 18005352]
27. Phillips CV. Publication bias in situ. BMC Med Res Methodol. 2004; 4:20. [PubMed: 15296515]
28. Siontis GC, Ioannidis JP. Risk factors and interventions with statistically significant tiny effects. Int J Epidemiol. 2011; 40(5):1292–1307. [PubMed: 21737403]
29. Ioannidis JP. Why most discovered true associations are inflated. Epidemiology. 2008; 19(5):640–648. [PubMed: 18633328]
30. Frøslie KF, Røislien J, Laake P, et al. Categorisation of continuous exposure variables revisited. A response to the Hyperglycaemia and Adverse Pregnancy Outcome (HAPO) Study. BMC Med Res Methodol. 2010; 10:103. [PubMed: 21062456]
31. Baik I, Ascherio A, Rimm EB, et al. Adiposity and mortality in men. Am J Epidemiol. 2000; 152(3):264–271. [PubMed: 10933273]
32. Schulgen G, Lausen B, Olsen JH, et al. Outcome-oriented cutpoints in analysis of quantitative exposures. Am J Epidemiol. 1994; 140(2):172–184. [PubMed: 8023805]
33. Altman DG, Lausen B, Sauerbrei W, et al. Dangers of using “optimal” cutpoints in the evaluation of prognostic factors. J Natl Cancer Inst. 1994; 86(11):829–835. [PubMed: 8182763]
34. Lin WY, Tsai SL, Albu JB, et al. Body mass index and all-cause mortality in a large Chinese cohort. CMAJ. 2011; 183(6):E329–E336. [PubMed: 21398246]
35. Durazo-Arvizu R, McGee D, Li Z, et al. Establishing the nadir of the body mass index–mortality relationship: a case study. J Am Stat Assoc. 1997; 92(440):1312–1319.
36. Durazo-Arvizu RA, Cooper RS. Issues related to modeling the body mass index–mortality association: the shape of the association and the effects of smoking status. Int J Obes (Lond). 2008; 32(suppl 3):S52–S55. [PubMed: 18695654]

37. Filardo G, Hamilton C, Hamman B, et al. Obesity and stroke after cardiac surgery: the impact of grouping body mass index. Ann Thorac Surg. 2007; 84(3):720–722. [PubMed: 17720366]

38. Filardo G, Hamilton C, Hamman B, et al. Categorizing BMI may lead to biased results in studies investigating in-hospital mortality after isolated CABG. J Clin Epidemiol. 2007; 60(11):1132–1139. [PubMed: 17938055]

39. Wong ES, Wang BC, Garrison LP, et al. Examining the BMI-mortality relationship using fractional polynomials. BMC Med Res Methodol. 2011; 11:175. [PubMed: 22204699]

40. Gilboa SM, Correa A, Alverson CJ. Use of spline regression in an analysis of maternal prepregnancy body mass index and adverse birth outcomes: Does it tell us more than we already know? Ann Epidemiol. 2008; 18(3):196–205. [PubMed: 18201903]

41. Janssen I, Mark AE. Elevated body mass index and mortality risk in the elderly. Obes Rev. 2007; 8(1):41–59. [PubMed: 17212795]

42. Chiolero A, Peytremann-Bridevaux I, Paccaud F. Associations between obesity and health conditions may be overestimated if self-reported body mass index is used. Obes Rev. 2007; 8(4):373–374. [PubMed: 17578386]

43. Faeh D, Roh L, Paccaud F, et al. Mortality risk of obesity and underweight is overestimated with self-reported body mass index. Epidemiology. 2014; 25(1):156–158. [PubMed: 24296932]

44. Spencer EA, Appleby PN, Davey GK, et al. Validity of self-reported height and weight in 4808 EPIC-Oxford participants. Public Health Nutr. 2002; 5(4):561–565. [PubMed: 12186665]

45. Plankey MW, Stevens J, Flegal KM, et al. Prediction equations do not eliminate systematic error in self-reported body mass index. Obes Res. 1997; 5(4):308–314. [PubMed: 9285836]

Table 1

BMI[a] Categories in the Articles Reviewed in the Current Study and Terminology Used in This Article

First Author, Year (Reference No.)	BMI Terminology and Category Definitions Used							
	Underweight	Low- Normal Weight	Mid- Normal Weight	High- Normal Weight	Low Overweight	High Overweight	Obesity Grade 1	Obesity Grades 2–3
Berrington de Gonzalez, 2010 (8)	15.0–18.4	18.5–19.9	20.0–22.4	22.5–24.9[b]	25.0–27.4	27.5–29.9	30.0–34.9	35.0–39.9 40.0–49.9
Calle, 1999 (9)	<18.5	18.5–20.4	20.5–21.9[c] and 22.0–23.4	23.5–24.9[b]	25.0–26.4	26.5–27.9[c] and 28.0–29.9	30.0–31.9[c] and 32.0–34.9	35.0–39.9 ≥40.0
Jacobs, 2010 (10)		18.5–<20	20.0–<22.5	22.5–<25[b]	25.0–<27.5	27.5–<30	30.0–<32.5[c] and 32.5–<35	≥35.0
Jee, 2006 (11)	<18.5	18.5–19.9	20.0–21.4[c] and 21.5–22.9	23.0–24.9[b]	25.0–26.4	26.5–27.9[c] and 28.0–29.9	30.0–31.9[c] and ≥32.0	
Joshy, 2014 (15)	15–18.49	18.5–19.99	20.0–22.49	22.5–24.99[b]	25.0–27.49	27.5–29.99	30.0–34.99	35–50
Klenk, 2009 (12)	<18.5	18.5–19.9	20.0–22.4	22.5–24.9[b]	25.0–27.4	27.5–29.9	30.0–34.9	≥35.0
Pischon, 2008 (13)	<18.5	18.5–<21.0	21.0–<23.5	23.5–<25.0[b]	25.0–<26.5	26.5–<28.0[c] and 28.0–<30.0	30.0–<35.0	≥35.0
Zheng, 2011 (14)	15.1–17.5	17.6–20.0	20.1–22.5	22.6–25.0[b]	25.1–27.5	27.6–30.0	30.1–32.5[c] and 32.6–35.0	35.1–50.0

Abbreviations: BMI, body mass index; NHLBI, National Heart, Lung and Blood Institute.

[a]BMI is calculated as height (kg)/height (m)^2. Standard BMI categories established by the National Heart, Lung, and Blood Institute (Bethesda, Maryland) and the World Health Organization (Geneva, Switzerland) are as follows: underweight, <18.5; normal weight, 18.5–24.9; overweight, 25.0–29.9, grade 1 obesity, 30.0–34.9, grade 2 obesity, 35.0–39.9, and grade 3 obesity, ≥40.0.

[b]Reference category.

[c]When there are 2 BMI subgroupings within a category, the lower set of values is used for the purposes of this paper.

Table 2

Summary and Study-Specific Adjusted[a] Hazard Ratios[b] for Men by Smoking Status and by Self-Reported Versus Measured Weight and Height Data

First Author, Year (Reference No.) by Type of Height and Weight Data	Underweight		Low- Normal Weight		Mid- Normal Weight		Low Overweight		High Overweight		Grade 1 Obesity	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Full Sample												
Measured height and weight												
Jee, 2006 (11)	1.35	1.30, 1.41	1.19	1.15, 1.23	1.16	1.13, 1.20	0.97	0.94, 1.00	0.99	0.95, 1.03	1.20	1.08, 1.34
Klenk, 2009 (12)	2.57	2.17, 3.05	1.44	1.25, 1.65	1.24	1.15, 1.34	1.01	0.95, 1.08	1.13	1.05, 1.21	1.25	1.15, 1.35
Pischon, 2008 (13)	2.30	1.84, 2.80	1.39	1.24, 1.57	1.03	0.94, 1.12	0.91	0.84, 0.98	0.96	0.88, 1.04	1.24	1.14, 1.35
Zheng, 2011[c] (14)	1.94	1.71, 2.20	1.40	1.28, 1.52	1.13	1.08, 1.18	0.97	0.93, 1.00	1.05	1.00, 1.11	1.21	1.14, 1.35
Zheng, 2011[d] (14)	1.63	1.39, 1.91	1.27	1.10, 1.47	1.10	0.95, 1.27	0.98	0.81, 1.18	0.99	0.75, 1.31	1.14	0.74, 1.75
Summary HR (measured)	1.89	1.45, 2.47[e]	1.33	1.20, 1.46[e]	1.14	1.09, 1.19	0.97	0.95, 0.99	1.03	0.97, 1.09[e]	1.23	1.18, 1.29
Self-reported height and weight												
Berrington de Gonzalez, 2010 (8)	1.85	1.72, 2.00	1.47	1.39, 1.56	1.15	1.12, 1.18	0.97	0.96, 0.99	1.05	1.02, 1.07	1.18	1.15, 1.21
Jacobs, 2010 (10)			1.64	1.40, 1.91	1.17	1.09, 1.26	0.95	0.90, 1.01	1.05	0.99, 1.12	1.05	0.96, 1.14
Summary HR (self-reported)	1.85	1.71, 2.00	1.52	1.38, 1.67	1.15	1.12, 1.18	0.97	0.95, 0.99	1.05	1.03, 1.07	1.12	1.00, 1.26[e]
Summary HR (all)	1.88	1.53, 2.31[e]	1.39	1.26, 1.53[e]	1.15	1.11, 1.18	0.97	0.96, 0.98	1.04	1.00, 1.07[e]	1.18	1.13, 1.24
Never Smokers Only												
Measured height and weight												
Jee, 2006 (11)	1.29	1.15, 1.44	1.13	1.04, 1.23	1.04	0.97, 1.12	0.99	0.93, 1.06	1.09	1.00, 1.19	1.54	1.27, 1.87
Klenk 2009 (12)	1.92	1.46, 2.53	1.28	1.04, 1.56	1.08	0.97, 1.20	1.03	0.95, 1.12	1.19	1.08, 1.31	1.35	1.21, 1.50
Pischon, 2008 (13)	1.19	0.54, 2.62	1.25	0.93, 1.67	0.97	0.80, 1.07	0.89	0.73, 1.07	1.05	0.86, 1.27	1.48	1.22, 1.79
Zheng 2011[c] (14)	1.82	1.44, 2.29	1.22	1.05, 1.42	0.99	0.93, 1.05	0.99	0.92, 1.07	1.19	1.06, 1.34	1.58	1.21, 2.07
Zheng 2011[d] (14)	1.55	1.22, 1.98	1.23	1.01, 1.52	1.06	0.87, 1.29	0.97	0.75, 1.24	1.01	0.71, 1.45	1.15	0.67, 1.98
Summary HR (measured)	1.57	1.29, 1.90[e]	1.18	1.10, 1.25	1.01	0.98, 1.05	0.99	0.95, 1.04	1.14	1.08, 1.20	1.42	1.31, 1.53
Self-reported height and weight												
Berrington de Gonzalez, 2010 (8)	1.37	1.09, 1.71	1.01	0.85, 1.20	1.00	0.93, 1.07	1.06	1.01, 1.12	1.21	1.14, 1.28	1.44	1.35, 1.53
Calle, 1999 (9)	1.26	1.02, 1.56	1.19	1.05, 1.34	1.09	1.00, 1.18	1.04	0.98, 1.10	1.09	1.02, 1.16	1.32	1.21, 1.45
Joshy, 2014 (15)	1.95	1.10, 3.44	1.89	1.24, 2.88	1.52	1.20, 1.92	1.02	0.82, 1.26	1.11	0.87, 1.42	1.56	1.21, 2.02
Summary HR (self-reported)	1.35	1.16, 1.57	1.22	0.97, 1.53	1.13	0.97, 1.31[e]	1.05	1.01, 1.09	1.14	1.05, 1.25	1.40	1.31, 1.51

First Author, Year (Reference No.) by Type of Height and Weight Data	Underweight		Low- Normal Weight		Mid- Normal Weight		Low Overweight		High Overweight		Grade 1 Obesity	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Summary HR (all never smokers)	1.49	1.31, 1.70	1.18	1.10, 1.27	1.04	0.99, 1.10[e]	1.02	0.99, 1.05	1.14	1.09, 1.19	1.41	1.35, 1.47

Abbreviations: CI, confidence interval; HR, hazard ratio.

[a]Adjustment factors are as follows: for Berrington de Gonzalez at al. (8), age, stratified by study, adjusted for alcohol, educational level, marital status, and physical activity; for Calle et al. (9), age, educational level, physical activity, alcohol use, marital status, use of aspirin, fat consumption, vegetable consumption, and (in women) use of estrogen replacement therapy; for Jacobs at al. (10), age, race, educational level, marital status, smoking, alcohol intake, height, and physical activity; for Jee at al. (11), age, alcohol intake, and participation in regular physical activity; for Joshy et al. (15), age (and excluding those who ever smoked at baseline, those with a history of cardiovascular disease at baseline, those with a history of cancer at baseline and during the first 2 years of follow-up); for Klenk et al. (12), age and smoking; for Pischon et al. (13), age, smoking, educational level, alcohol consumption, physical activity, and height; and for Zheng et al. (14), age, sex, educational level, urban/rural residence, marital status, and baseline comorbidities (and excluding subjects with less than 3 years of follow-up).

[b]Reference category is men of high-normal weight.

[c]Population consisted of East Asians.

[d]Population consisted of Indians and Bangladeshis.

[e]Statistically significant heterogeneity ($P < 0.05$).

Table 3

Summary and Study-Specific Adjusted[a] Hazard Ratios[b] for Women by Smoking Status and by Self-Reported Versus Measured Weight and Height Data

First Author, Year (Reference No.) by Type of Height and Weight Data	Underweight		Low- Normal Weight		Mid- Normal Weight		Low Overweight		High Overweight		Grade 1 Obesity	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Full Sample												
Measured height and weight												
Jee, 2006 (11)	1.16	1.09, 1.23	1.11	1.05, 1.17	0.99	0.95, 1.04	0.98	0.94, 1.03	1.02	0.97, 1.08	1.16	1.06, 1.28
Klenk, 2009 (12)	1.40	1.21, 1.62	1.11	0.99, 1.26	1.06	0.98, 1.14	1.05	0.98, 1.13	1.10	1.02, 1.19	1.29	1.20, 1.40
Pischon, 2008 (13)	1.71	1.44, 2.01	1.22	1.00, 1.34	1.00	0.92, 1.09	1.01	0.92, 1.11	1.07	0.97, 1.18	1.17	1.07, 1.29
Zheng, 2011[c] (14)	1.74	1.53, 1.98	1.28	1.18, 1.39	1.03	0.98, 1.09	1.00	0.95, 1.06	1.07	1.00, 1.15	1.19	1.06, 1.34
Zheng, 2011[d] (14)	1.55	1.25, 1.92	1.23	1.01, 1.51	1.08	0.88, 1.32	0.98	0.77, 1.25	0.89	0.65, 1.22	0.97	0.64, 1.47
Summary HR (measured)	1.49	1.22, 1.81[e]	1.18	1.10, 1.27	1.02	0.99, 1.05	1.00	0.97, 1.03	1.05	1.02, 1.09	1.20	1.14, 1.27
Self-reported height and weight												
Berrington de Gonzalez, 2010 (8)	1.78	1.70, 1.85	1.25	1.21, 1.30	1.03	1.01, 1.06	1.05	1.03, 1.07	1.14	1.11, 1.17	1.31	1.28, 1.34
Jacobs, 2010 (10)			1.34	1.19, 1.51	1.17	1.08, 1.27	1.06	0.97, 1.15	1.08	0.98, 1.19	1.12	1.00, 1.26
Summary HR (self-reported)	1.78	1.71, 1.85	1.26	1.20, 1.33	1.09	0.96, 1.24[e]	1.05	1.03, 1.07	1.13	1.09, 1.17	1.22	1.05, 1.43[e]
Summary HR (all)	1.54	1.26, 1.88[e]	1.21	1.15, 1.28[e]	1.04	1.00, 1.08	1.03	1.00, 1.05	1.08	1.03, 1.13[e]	1.21	1.14, 1.29[e]
Never Smokers Only												
Measured height and weight												
Jee, 2006 (11)	1.17	1.09, 1.26	1.08	1.02, 1.15	1.01	0.96, 1.06	0.99	0.94, 1.05	1.05	0.99, 1.11	1.18	1.07, 1.31
Klenk, 2009 (12)	1.29	1.09, 1.52	1.11	0.97, 1.27	1.01	0.93, 1.10	1.05	0.98, 1.13	1.09	1.01, 1.19	1.29	1.19, 1.40
Pischon, 2008 (13)	1.44	1.08, 1.92	1.09	0.93, 1.27	1.00	0.88, 1.13	1.00	0.87, 1.15	1.12	0.98, 1.30	1.25	1.09, 1.43
Zheng, 2011[c] (14)	1.72	1.53, 1.94	1.26	1.16, 1.37	1.02	0.97, 1.09	1.00	0.94, 1.07	1.08	1.01, 1.15	1.18	1.05, 1.33
Zheng, 2011[d] (14)	1.54	1.24, 1.92	1.25	1.02, 1.53	1.08	0.88, 1.32	0.98	0.77, 1.26	0.90	0.66, 1.23	0.97	0.64, 1.48
Summary HR (measured)	1.41	1.17, 1.70[e]	1.15	1.06, 1.24	1.01	0.98, 1.05	1.01	0.97, 1.04	1.07	1.03, 1.11	1.23	1.17, 1.30
Self-reported height and weight												
Berrington de Gonzalez, 2010 (8)	1.47	1.33, 1.62	1.14	1.07, 1.22	1.00	0.96, 1.04	1.09	1.05, 1.14	1.19	1.14, 1.24	1.44	1.38, 1.50
Calle, 1999 (9)	1.36	1.25, 1.48	1.10	1.04, 1.16	1.00	0.95, 1.05	1.07	1.01, 1.13	1.10	1.04, 1.17	1.30	1.22, 1.39
Joshy, 2014 (15)	1.97	1.38, 2.81	1.72	1.28, 2.32	1.43	1.14, 1.80	1.01	0.80, 1.28	1.09	0.84, 1.42	1.27	0.98, 1.66
Summary HR (self-reported)	1.45	1.30, 1.63	1.17	1.05, 1.31[d]	1.04	0.96, 1.13[e]	1.08	1.04, 1.12	1.15	1.07, 1.22	1.36	1.25, 1.49

First Author, Year (Reference No.) by Type of Height and Weight Data	Underweight		Low- Normal Weight		Mid- Normal Weight		Low Overweight		High Overweight		Grade 1 Obesity	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Summary HR (all never smokers)	1.44	1.29, 1.61[e]	1.15	1.09, 1.22[d]	1.01	0.98, 1.04	1.04	1.01, 1.07	1.10	1.05, 1.15[e]	1.27	1.19, 1.37[e]

Abbreviations: CI, confidence interval; HR, hazard ratio.

[a] Adjustment factors are as follows: for Berrington de Gonzalez at al. (8), age, stratified by study, adjusted for alcohol, educational level, marital status, and physical activity; for Calle et al. (9), age, educational level, physical activity, alcohol use, marital status, use of aspirin, fat consumption, vegetable consumption, and (in women) use of estrogen replacement therapy; for Jacobs at al. (10), age, race, educational level, marital status, smoking, alcohol intake, height, and physical activity; for Jee at al. (11), age, alcohol intake, and participation in regular physical activity; for Joshy et al. (15), age (and excluding those who ever smoked at baseline, those with a history of cardiovascular disease at baseline, those with a history of cancer at baseline and during the first 2 years of follow-up); for Klenk et al. (12), age and smoking; for Pischon et al. (13), age, smoking, educational level, alcohol consumption, physical activity, and height; and for Zheng et al. (14), age, sex, educational level, urban/rural residence, marital status, and baseline comorbidities (and excluding subjects with less than 3 years of follow-up).

[b] Reference group is women of high-normal weight.

[c] Population consisted of East Asians.

[d] Population consisted of Indians and Bangladeshis.

[e] Statistically significant heterogeneity ($P < 0.05$).

Table 4

NHANES Example[a] of the Effect on Hazard Ratios of Varying the BMI[b] Reference Category

Sex and BMI Category	BMI Reference Category	Hazard Ratio		
		Overweight (BMI 25.0–29.9)	Grade 1 Obesity (BMI 30.0–34.9)	Grades 2–3 Obesity (BMI ≥35.0)
Men				
<25.0	Normal and underweight	0.95	1.16[c]	1.77[c]
18.5–24.9	Normal weight	0.97	1.19[c]	1.82[c]
20.0–24.9	Mid- and high-normal weight	0.99	1.21[c]	1.85[c]
20.0–22.9	Mid-normal weight	0.90	1.10	1.68[c]
23.0–24.9	High-normal weight	1.05	1.28[c]	1.96[c]
Women				
<25.0	Normal and underweight	0.98	1.22[c]	1.67[c]
18.5–24.9	Normal weight	1.02	1.28[c]	1.75[c]
20.0–24.9	Mid- and high-normal weight	1.03	1.28[c]	1.76[c]
20.0–22.9	Mid-normal weight	1.00	1.24[c]	1.70[c]
23.0–24.9	High-normal weight	1.06	1.32[c]	1.81[c]

[a]Based on example from NHANES I-II-III with no more than 25 years of follow-up and subjects less than 70 years of age at baseline.

[b]Weight (kg)/height (m)^2.

[c]Significantly different from 1 ($P < 0.05$).